M1-086E
C.2.

A TESTING PROTOCOL FOR AN INTERNATIONAL CLASSIFICATION OF DISEASES (ICD)

Karol P. Krotki
David G. Paton
Institutional and Agriculture
Survey Methods Division
Statistics Canada

August 1983

TABLE OF CONTENTS

A. Introduction

B. Need for a Revised ICD

C. General Testing Considerations

D. Desirable Properties of an ICD

D.1 Needs of analysts

D.2 Collection and coding

D.3 Flexibility

D.4 Cost considerations

Bibliography

A. Introduction

In the introduction to the report of the Conference on Health Statistics for the Year 2000 the authors state:

> "All health management, including planning, monitoring and evaluation, must be underpinned by health information, much of which will be in statistical form. Without this, resources allocated to the health field might be seriously misdirected.
>
> The form of the health information system in any particular country will depend upon the latter's state of development and the characteristics of its health services. The conference discussed a range of methods of acquiring health data including lay reporting, household surveys and other techniques relevant to morbidity, mortality, hospital care and primary care and other health determinants. There is a need for various forms of reporting of health problems, not all of which may be relevant in each country.
>
> An essential component of all statistical systems is classification. A common thread running throughout the conference was the various types of classifications and their inter-relationships required by modern statistical practices. This is a particularly important topic when WHO will shortly be formulating its policy for the 10th Revision of the ICD."

This report is an attempt to justify and define a testing protocol for a proposed revision to the ICD-9. Given the cost of introducing a new system, a thorough testing and evaluation program is necessary. Hopefully this testing protocol is a first positive step in such a direction. Emphasis in the report is placed on establishing a testing protocol rather than elaborating detailed testing strategies. The goal of the report is to outline in very general terms what aspects of an ICD should be tested, why these aspects should be tested, and how the testing should be done. It is only in answer to the third question that this report delves into the area of testing strategy and methodology. Such testing strategies will be discussed and proposed but only in general terms and details will be avoided.

The general nature of this report also stems from the authors' instructions to establish as universal a protocol as possible. The guidelines and recommendations should be relevant to, understood by, and practicable for a wide audience both nationally and internationally.

The next section of this report (Section B) discusses the current version of the ICD (ICD-9) and its perceived shortcomings. A few introductory and general comments on testing an ICD are presented in

Section C which is followed by a list of desirable properties of an ICD (Section D). In the final section (E), actual testing strategies are proposed to evaluate an ICD along the dimensions outlined in section D. The bibliography offers a random selection of related articles and documents that might prove useful as further reading.

B. Need for a Revised ICD

It is revealing to review notes from recent WHO classification meetings. There seems to be an overwhelming sentiment in favor of a change (in some cases "a radical change") of the classification system. The "appropriateness of the ICD-9" is often questioned in a variety of general areas. Allusions are made to a "multi-dimensional approach" and suggestions are made for "different ways of structuring" and for a "better balance between chapters".

Some comments refer to the balancing act that an ideal ICD has to perform. One such act is the balance "between the need to keep up-to-date and the desire to change very little". Another is the "need for broad classification (for international comparisons) together with specificity (for diagnostic indexing and epidemiological research)". The implications in both of the above examples is that the ICD-9 is not as good as it should, or could, be. Finally the phrase "more flexibility" crops up repeatedly and indicates a need for some change in this direction.

There is a feeling shared by many users that changes are needed. The shortcomings in the existing systems have been loudly articulated but unfortunately not very precisely. The questions that do not seem to have been widely addressed are precisely what changes are needed and how should they be implemented.

Nevertheless, some particular shortcomings of the ICD-9 have been articulated in more specific terms. For example, it has been pointed out that it is not easy to add new codes to the ICD-9. Because of the structure of the index (4 digits long) in general there are only a limited number of free codes. In particular, there may be very few or none at the end of a chapter to which it is desired to add a new code. In response to this problem, a fifth digit has been added in places and supplementary codes have been resorted to (external causes (E codes), supplementary V codes, and morphology codes).

Other areas of improvement to the ICD-9 that have been precisely specified are: ability to identify late effects, overlap between alcoholism and physical effects, ability to identify drug reactions, and the psychiatric section.

Another specifically identified problem of the ICD-9 is that it may be difficult to equate it to other classfication systems. The ICD-9 also does not lend itself easily to manipulation such as various forms of collapsing to create desired aggregations of codes. By the same token, manipulation of codes to create new schema, hierarchies, and structures is not easily done.

C. General Testing Considerations

As mentioned in the introduction, one goal of this report is to establish a testing protocol that is not dependent on the source, type, and environment of the data. The system to be tested is not only supposed to be an International Classification System (and, as such, applicable to all countries) but also must satisfy the criterion that it be universally acceptable along many other dimensions.

One of the desirable traits of an ICD system is the ease with which it can serve as a translating medium between written diagnoses and coded data to be used for analysis. The testing protocol to be established here will be based on the assumption that these written diagnoses are the fixed input which is not to be evaluated. The testing is to center on the ease with which coders can translate these written diagnoses into coded data.

We suggest that the most recent version of the ICD be used as a standard of comparison since the replacement version should be at least as good as the existing version.

D. Desirable Properties of an ICD

Any ICD should strive to satisfy a number of desirable criteria. For ease of presentation, these have been arranged into four categories with detailed requirements listed in each.

D.1 Needs of analysts: An ICD must satisfy the needs of its users (primarily health managers) by allowing, and hopefully expediting, the analyses required.

1. An ICD should cover all diseases and disease concepts.

2. An ICD should allow for meaningful comparisons with statisticscollected under previous ICD's.

3. An ICD shoud facilitate aggregation of codes into meaningful categories, including both internally defined categories for the purpose of statistical presentation and study, and externally defined categories to allow comparisons with other studies.

4. An ICD should be organized in such a way that analysts can easilydetermine the set of codes that match a set of diagnoses.

5. An ICD should be as universal as possible to permit analyses in a wide variety of environments. As much as possible, the ICD should be independent of factors such as: language and nationality of users; data source; skill, experience, and training of users; level of sophistication of computer system.

D.2 Collection and coding: An ICD must be compatible with the various ways that health data may be collected and coded.

6. An ICD should give similar results when coded in different types of hospitals, in different countries, or by coders with different backgrounds. In other words, coder variance and bias should be as small as possible.

7. An ICD coding system should be easy to learn, with uncomplicated and unambiguous rules.

8. An ICD coding system should be efficient to use (i.e., given a diagnosis, the correct code can be quickly determined).

9. An ICD should include codes from data reported by non-professionals.

10. An ICD should include codes requiring a minimum of administration and processing, highlighting the concerns of developing countries.

D.3 Flexibility: An ICD should be flexible enough to both serve as the core linking a family of more specialized classifications, and to extend its life by allowing new categories to be added to its structure.

11. An ICD should allow the easy addition of new codes for new disease concepts or classes.

12. An ICD should serve as the keystone in a set of related classifications, thus integrating the nosographies for: household surveys, reasons for health care contact, diseases, etc.

D.4 Cost considerations: The benefits of a revised ICD have to be weighed against the costs associated with its introduction and use. These costs include:

13. Cost of retraining coders.

14. Cost of rewriting computer software (e.g., computer storage and retrieval processes).

15. Cost of running the computer system (coding, collapsing, storage, retrieval, and dissemination).

16. Cost of redesigning dissemination procedures.

17. Cost of developing comparisons with data based on the ICD-9.

E. Testing Strategies

Some of the criteria identified above lend themselves better to review by subject-matter experts or ICD users than they do to quantitative testing. For the first class of criteria, subject-matter specialists and ICD users should be consulted and asked to review the classification following suggested guidelines.

Criteria 1, 3, 4, 5, 9, 10, 11, 12, 13, 16 and 17 are in this class. The others (2, 6, 7, 8, 14, 15) can be tested using statistical or computer analysis. In these cases known techniques of experimental design, sampling, applied mathematics, and data processing, using both real and manufactured data and using both the old and the new ICD should be used. The tests should be carried out in several environments and under a variety of collection modes, and caution should be exercised in generalizing the results to others.

1. Does the ICD cover all diseases and disease concepts?

   This question should be investigated by requesting a group of subject-matter experts to review the ICD. It is recommended that this task be delegated to a large number of experts from various specialities.

2. Does the ICD allow for meaningful comparisons with statistics collected under previous ICD's?

   This is not a question suited to quantitative analysis but, because of the large number of codes which must be cross-referenced, analysis of this requirment should be computer assisted. Conversion tables to previous ICD's should be put into machine-readable form and mappings in both directions determined.

   If the mappings are "one-to-one" then the new code is the same as the old one in terms of specificity and all comparisons are possible. However, it is unlikely that the mapping will be one-to-one. The new code may be more specific than the old code, with respect to a class of diseases in which case one old code will correspond to more than one new code. The converse, with one new code corresponding to several old codes may also occur. These conditions place restrictions on the collapsibility of the codes. Analysis to identify groups of interconnected codes can identify comparison problems.

3. Does the ICD facilitate aggregation of codes into meaningful categories?

   This question should be considered with the help of subject-matter experts and users. Respondents should be asked to suggest meaningful categories or collapsing schemes; these may correspond to categories of their own device or to categories used in their specialities. The ease and possibility of collapsing the ICD to

match these categories can then be investigated both informally and by attempting to write programs to perform the collapsing. Conversely, the categories that result from simply ignoring the final digit(s) of the codes should be given to the subject-matter experts for their comments on whether these most simple collapsing schemes lead to meaningful categories.

4. Can analysts easily determine which codes match a given diagnosis?

   The answer to this question will depend to a large extent on the quality of the documentation, indexes, and cross-references associated with the ICD, as well as on its logical structure. The opinions of those involved in developing the ICD and its accompanying reference material should be solicited. Additionally, analysts should be given the documentation and asked to determine and comment on the ease with which they can find the codes that correspond to given diagnoses. They should be asked to compare the ICD to the previous verson especially in terms of strengths and weaknesses in the documentation and in terms of the effect of its logical structure on the ease of finding codes.

5. Is analysis possible in a wide variety of environments?

   Various agencies and individuals which use the ICD codes should be polled to ensure that no characteristics of the new code make it incompatible with use in their environmnt, say, due to special characters, alphabetic charcters, length of codes, variability of length of codes, etc.

6. Does the ICD give similar results when applied in different environments?
7. Can coders be easily trained?
8. Is the coding system efficient to use?

   These questions must be investigated statistically with carefully planned and executed experiments. For all these questions, the ICD must be compared to a previous version and groups of coders with none or very litle experience with either verison will be needed. Some coders will be trained in both versions, and the results of their coding compared.

   For the experiment addressing itself to Requirement 6 the comparison between the two verisons is not as important as the comparison between coders of different backgrounds (physicians, trained coders, lay reporters), and between different levels of any other environmental factor that might affect reporting (private or public hospital, etc.)

   For Requirement 7, we are interested in the speed with which coders become proficient, so the two training programmes should progress in parallel as much as possible, and at several stages in the programme the coders should be tested.

For Requirement 8, a variety of sources of diagnoses should be used to compare not only the relative efficiencies of the two versions, but also to investigate the effect of the source of diagnoses on the efficiencies.

In these requirements, the basic strategy will use as input data the diagnoses to be coded by the agents. If the replication method is adopted each coder will code the same input twice, using two versions of the ICD. If the "contamination" risk is too high, an interpenetrated model will be used in which a random half of the agents will use one version and the other half will use the second version. The ultimate objective is to measure the effect of the "treatment" (the classification system) on the coding process.

Is there a significant difference between the two systems and, if so, what is the size of this difference? Care will have to be taken in designing these experiments since it is unlikely that there will be a large number of coders to participate in the study.

9. Does the coding system include codes for data reported by non-professionals?

   This question should be considered by the ICD designer and by those agencies using lay reporters to collect and code data under the ICD system.

10. Does the coding system include codes highlighting the concerns of developing nations?

    This is another question that should be answered by the agencies and users involved. The codes should be compatible with the needs of the user nations and with the need to make comparisons with data based on more detailed ICD codes.

11. Does the ICD allow for the addition of new codes?

    This can be best answered by the ICD's designers. Subject-matter experts shold be asked to suggest the most likely areas in which new codes will be needed. The ability to add new codes in the corresponding sections of the ICD should then be verified.

12. Will the ICD integrate nosographies for a variety of uses?

    This will only happen if the variety of users can be convinced of its utility. The ICD's designers should develop or encourage the development of related classification for special uses, and must consider whether the reasons for other users to link their nosographies to the ICD are compelling enough.

13. Cost of retraining coders.

    Training experts should be able to provide estimates of this cost. More precise figures can be obtained by actually monitoring the retraining process. However, it is recommended that this not be emphasized since intangible costs cannot be measured and since measuring even direct costs is a difficult task.

14. Cost of rewriting computer software.

    A revised ICD will require a change to computer software. Programs of retrieval, storage, analysis, and dissemination will have to be adjusted. Presumably the cost will depend on the nature of the software at any particular installation. However, it should be possible to make a general assessment of the effect of a revised ICD on rewriting computer software. In addition, several examples of computer systems may be studied with a view to deciding how they would have to be changed.

15. Cost of running the computer system.

    A similar approach to the one proposed for requirement 14 is recommended here. A general assessment should be coupled with an analysis of a few particular cases. The emphasis here is on the performance (i.e., demand on computer time and storage space) of the computer software under the revised ICD

16. Cost of redesigning dissemination procedures.

    Technical experts, familiar with dissemination procedures used for the ICD-9, should analyze the extent to which the change to the ICD will affect dissemination procedures. Possible areas of concern are structure of tables, format of reports, and methods of analyzing and presenting data.

17. Cost of developing comparisons with data based on the ICD-9.

    It is possible (and recommended) to simultaneously evaluate a number of feasible alternatives to a revised ICD. In this way, the costs of evaluation may be reduced and a larger number of potential revisions can be tested.

BIBLIOGRAPHY

Cook, D., et al., "The accuracy and completeness of neonantal mortality statistics for congenital heart disease, Canadian Journal of Public Health, 73 (Nov./Dec. 1982): 420-23.

Densen, P., E. Balawith, and N. Deerdorff, "Medical care plans as a source of morbidity data", Milbank Memorial Fund Quarterly.

Gittelsohn, A. and J. Sennag, "Studies on the reliability of vital and health records: 1. Comparison of cause of death and hospital record diagnoses", American Journal of Public Health, 69(7) (July 1979): 680-889.

Israel, R. and D. Green, "A trial of an alpha-numeric code structure for ICD", Meeting of Heads of Who Centres for Classficiation of Diseases, Moscow, USSR, 18-22 June 1979 (ICD/C/79.10).

Krotki, K., "Testing protocol for a proposed revision to the ICD-9" (Discussion Paper), IASMD, Statitics Canada, September 1982.

Lubits, J., "Different data systems, different conclusions? Comparing hospital use data for the aged from four data systems", Health Care Financing Review, Spring 1981, pp. 41-60.

National Center for Health Statistics, "Estimates of selected comparability ratios based on dual coding of 1976 death certificates by the eighth and ninth revisions of the International Classification of Diseases", Monthly Vital Statistics Report 28(11) (29 Geb. 1980).

Nosology Reference Centre, "A biaxial concept for disease classification", Health Division, Statistics Canada, June 1983.

Office of Population Censusses and Surveys, "Cohort studies: new developments", Studies on Medical and Population Subjects No. 25, London: Her Majesty's Stationery Office, 1973.

Paton, D., "International Classification of Diseases - Testing protocol" (Discussion paper), IASMD, Statistics Canada, July 1983.

Slee, V., "Clinical entity coding", personal communication, 1983.

Tenney, J., "Daignsotic precision for insurance records: a physicians survey", Inquiry 5(4) (1968): 14-18. Institute of medicine, Reliability of National Hospital Discharge Survey Data Study, Washington: National Academy of Services, 1980.

Thurlbeck, W., "Accuracy of clinical diagnosis in a Canadian Teaching hospital", Canadian Medical Association Journal, 125 (Sept. 1981): 443-447.

World Health Organisation, "Evaluation of the ninth revision of the International Classification of Diseases", Report on a WHO Workshop, banterburg 1-5 September 1981.

World Health Organization, "Report of the Bibliography Conference on Health Statistics for the year 2000", 27 September - 1 October 1982, Bellagio.

World Health Organization, "Evaluation of the ninth revision of the International Classification of Diseases", Report on a WHO Workshop, Canterburg 1-5 September 1981.

M-086F
c. 2.

PROTOCOLE D'ESSAI D'UNE CLASSIFICATION INTERNATIONALE DES MALADIES (CIM)

Karol P. Krotki
David G. Paton
Division des méthodes
d'enquête-institutions
et agriculture
Statistique Canada

Août 1983

TABLE DES MATIÈRES

A. Introduction

B. Besoin d'une CIM révisée

C. Considérations générales sur l'essai

D. Caractéristiques souhaitables d'une CIM

D.1 Besoins des analyses

D.2 Collecte et codage

D.3 Souplesse

D.4 Considérations sur les coûts

Bibliographie

A. Introduction

Dans l'introduction du rapport de la conférence sur la statistique de la santé pour l'an 2000, les auteurs déclarent:

> "Toutes les fonctions de gestion de la santé, y compris la planification, la surveillance et l'évaluation, doivent reposer sur des renseignements dont la plupart seront d'ordre statistique. Sans ces renseignements, l'attribution des ressources au secteur de la santé pourrait être très mal effectuée.
>
> Les structures de l'appareil de données sur la santé dans un pays en particulier dépendent du stade de développement de ce dernier et des caractéristiques de ses services de santé. La conférence a permis de discuter d'un éventail de méthodes d'acquisition de données sur la santé, y compris des déclarations par des non-spécialistes, des enquêtes-ménages et d'autres techniques pertinentes à la morbidité, à la mortalité, aux soins hospitaliers et primaires ainsi que d'autres déterminants de la santé. La déclaration des problèmes de santé a besoin d'être faite sous diverses formes et celles-ci ne se retrouvent pas toutes dans chaque pays.
>
> Tous les appareils statistiques comportent un élément essentiel: une classification. La conférence reposait d'ailleurs sur les divers types de classification et sur les interrelations qu'elles doivent présenter aux fins des méthodes statistiques modernes. Il s'agit là d'une question particulièrement importante dont l'OMS devra tenir compte lorsqu'elle formulera bientôt ses directives pour la dixième révision de la CIM."

Le présent rapport vise à justifier et à définir un protocole d'essai en vue d'une révision proposée de la CIM-9. Vu le coût d'introduction d'un nouveau système, il est nécessaire de se doter d'un programme d'évaluation et d'essai en profondeur. Il est à espérer que ce protocole d'essai sera un premier pas dans cette direction. Nous insistons d'ailleurs sur un protocole d'essai plutôt que sur des stratégies détaillées d'essai. Le but du rapport est de présenter en des termes très généraux les aspects d'une CIM qui devraient être mis à l'essai, la raison derrière chaque aspect et la façon d'effectuer les essais. Si le présent rapport touche à une méthodologie et à une stratégie d'essai, ce n'est que pour répondre à la troisième question. Nous ne nous en tiendrons donc qu'à une discussion générale des stratégies d'essai.

La nature générale du rapport découle aussi du mandat qui avait été confié aux auteurs, c'est-à-dire établir un protocole aussi universel que possible. Les lignes directrices et les recommandations devront donc toucher le plus large auditoire possible à l'échelle nationale et internationale.

A la section B du rapport, nous traiterons de la version actuelle de la CIM-9 et de ses lacunes évidentes. A la section C, nous illustrerons par quelques remarques générales la façon d'éprouver une CIM puis nous dresserons une liste des caractéristiques souhaitables propres à une classification des maladies

(section D). Enfin, dans la section E, nous proposerons des stratégies réelles pour évaluer une CIM selon les caractéristiques énumérées à la section D. La bibliographie présente un choix non exhaustif d'articles et de documents connexes pouvant s'avérer utile aux lecteurs qui désirent en savoir davantage sur le sujet.

B. Besoin d'une CIM révisée

La lecture des comptes rendus de récentes réunions de l'OMS sur la classification est très instructive. Il semble y avoir un sentiment général en faveur d'une révision (dans certains cas "un changement radical") du système de classification. On remet souvent en question, dans divers domaines généraux, le caractère approprié de la CIM-9. On fait également allusion à une approche multidimensionnelle et on propose différents moyens de structure et un meilleur équilibre entre les chapitres.

Certaines observations ont trait à l'équilibre qu'une CIM idéale doit chercher à atteindre. En effet, il est question d'obtenir un équilibre entre le besoin de se tenir à jour et le désir de changer le moins possible. On fait également mention du besoin d'une classification étendue (à des fins de comparaison internationale) ainsi que d'une spécificité (pour le codage des diagnostics et la recherche épidémiologique). Ces deux exemples suffisent à démontrer que la CIM-9 n'est pas aussi bonne qu'elle pourrait ou devrait l'être. En dernier lieu, on retrouve souvent les termes "souplesse accrue" ce qui indique un besoin de changement en ce sens.

Bon nombre d'utilisateurs sont d'avis que des changements s'imposent. Les lacunes des systèmes actuels ont été largement critiquées mais, malheureusement, de façon très imprécise. Ce qui ne semble pas avoir été abordé en général, ce sont précisément le genre de changements et la façon dont ils devraient être mis en oeuvre.

Quoi qu'il en soit, quelques lacunes précises de la CIM-9 ont été cernées de façon précise. Ainsi, on a fait remarquer qu'il n'était pas facile d'ajouter de nouveaux codes à la CIM-9. En raison de la structure de l'index (à quatre chiffres) en général, il n'y a qu'un nombre limité de codes non utilisés. De façon particulière, un chapitre pourrait bien ne comporter à peu près aucun code lorsqu'on désire en ajouter un nouveau. Pour faire face à ce problème, un cinquième chiffre était ajouté et des codes supplémentaires ont été utilisés (causes externes (codes E), codes V supplémentaire, et codes de morphologie).

La CIM-9 a également fait l'objet d'autres critiques précises: incapacité de cerner les répercussions tardives, chevauchement entre les effets de l'alcoolisme et les effets physiques, incapacité de cerner les réactions aux drogues et lacunes dans la section des problèmes psychiatriques.

Les participants ont également précisé un autre problème de la CIM-9, soit le rapprochement avec d'autres systèmes de classification. En effet, la CIM-9 ne se prête pas facilement à une manipulation, par exemple à diverses formes de regroupement afin de créer des agrégations de codes. Du même coût, toute manipulation de codes dans le but de créer de nouveaux schémas, de nouvelles hiérarchies ou structures n'est pas chose facile.

C. Considérations générales relatives aux essais

Comme il est mentionné dans l'introduction, le présent rapport vise entres autres à établir un protocole d'essai qui ne soit pas dépendant de la source, du type et du contexte des données. Le système à éprouver n'est pas censé être uniquement une classification internationale des maladies (et, comme tel, applicable à tous les pays), mais il doit aussi être accepté universellement sous bon nombre d'autres aspects.

Une des caractéristiques souhaitables d'une CIM est la facilité avec laquelle elle peut servir d'intermédiaire entre les diagnostics écrits et les données codées servant à l'analyse. Le protocole d'essai que nous tenterons d'établir ici sera basé sur l'hypothèse voulant que ces diagnostics écrits constituent la matière première immuable qui n'a pas besoin d'être évaluée. L'essai se bornera à évaluer la facilité avec laquelle les codeurs peuvent convertir ces diagnostics écrits en codes.

Nous proposons donc que la plus récente version de la CIM serve de norme de comparaison vu que sa version de remplacement devrait être aussi bonne que la version actuelle.

D. Caractéristiques souhaitables d'une CIM

Toute CIM doit viser à respecter un certain nombre de critères souhaitables. Pour faciliter la présentation, ces critères ont été disposés en quatre catégories, chacune présentant des exigences détaillées.

D.1 Besoins des analystes: un CIM doit répondre au besoin de ses utilisateurs (surtout les gestionnaires de la santé) en permettant, et au mieux en accélérant, les analyses nécessaires.

1. Une CIM doit porter sur toutes les maladies et tous les concepts liés aux maladies.

2. Une CIM doit permettre des comparaisons utiles avec les statistiques recueillies aux termes des CIM antérieures.

3. Une CIM doit faciliter l'agrégation de codes en catégories valables, y compris les catégories internes aux fins de la présentation et de l'étude de statistiques, et des catégories externes permettant la comparaison avec d'autres études.

4. Une CIM doit être structurée de façon que les analystes puissent facilement déterminer l'ensemble de codes qui correspond à un ensemble de diagnostics.

5. Une CIM doit être aussi universelle que possible afin de permettre la réalisation d'analyses dans divers milieux. Dans la mesure du possible, la CIM devra être étrangère aux facteurs suivants: langue et nationalité des utilisateurs; source des données; aptitudes, expérience et formation des utilisateurs; niveau de complexité du système informatique.

D.2 Collecte et codage: Une CIM doit pouvoir se plier aux diverses méthodes de collecte et de codage des données sur la santé.

6. Une CIM doit donner des résultats similaires, peu importe que le codage soit fait dans des hôpitaux différents, dans des pays différents ou par des codeurs possédant une formation différente. En d'autres mots, la variance et le biais découlant des codeurs devraient peu porter à conséquence.

7. Le système de codage d'une CIM devrait être facile à comprendre et ne pas comporter de règle ambigüe ou compliquée.

8. Le système de codage d'une CIM devrait être facile à utiliser (c'est-à-dire qu'il devrait être facile d'établir le code d'un diagnostic donné).

9. Une CIM devrait comporter des codes pour tenir compte des données déclarées par des non-spécialistes.

10. Une CIM devrait comporter des codes ne nécessitant qu'un minimum d'administration et de traitement tout en faisant ressortir les préoccupations des pays en développement.

D.3 Souplesse: Une CIM devrait comporter suffisamment de souplesse pour servir de base au rapprochement d'une famille de classifications très spécialisées et pour prolonger sa durée en permettant l'ajout de nouvelles catégories à sa structure fondamentale.

11. Il devrait être facile d'ajouter à une CIM de nouveaux codes pour tenir compte de nouvelles catégories ou concepts de maladie.

12. Une CIM devrait servir de pierre d'achoppement dans un ensemble de classifications connexes, intégrant de ce fait les nosographies pour: les enquêtes-ménages, les raisons des contacts en vue des soins de santé, les maladies, etc.

D.4 Considérations sur les coûts: Il faudra évaluer les avantages d'une CIM révisée en fonction des coûts qu'entraîneront son introduction et son utilisation. Parmi ces coûts, mentionnons:

13. Coût de recyclage des codeurs.

14. Coût des nouveaux logiciels informatiques (par exemple processus de stockage et de documentation informatiques).

15. Coût du passage des programmes (codage, regroupement, stockage, documentation et diffusion).

16. Coût de reconception des procédures de diffusion.

17. Coût des comparaisons avec les données basées sur la CIM-9.

E. Stratégies d'essai

Quelques-uns des critères susmentionnés se prêtent mieux à un examen par les spécialistes ou les utilisateurs de la CIM qu'à un essai quantitatif. Dans le cas de la première catégorie de critères, il faudra consulter les spécialistes et les utilisateurs de la CIM et leur demander d'examiner la classification selon les lignes directrices proposées.

Les critères 1, 3, 4, 5, 9, 10, 11, 12, 13, 16 et 17 appartiennent à cette catégorie. Les autres (2, 6, 7, 8, 14, 15) peuvent être mis à l'essai dans une analyse informatique ou statistique. Dans ces cas, il faudrait utiliser les techniques connues de conception expérimentale, d'échantillonnage, de mathématiques appliquées et de traitement de données en se servant à la fois de données réelles et de données simulées, ainsi que de la version ancienne et nouvelle de la CIM. Les tests devront être effectués à plusieurs endroits et avec diverses méthodes de collecte; il faudra également prendre garde de généraliser les résultats.

1. La CIM porte-t-elle sur toutes les maladies et les concepts de maladie?

   Il faudra demander à un groupe de spécialistes d'examiner cette question. On recommande que cette tâche soit déléguée à un grand nombre d'experts de spécialités diverses.

2. La CIM se prête-t-elle à des comparaisons valables avec d'autres statistiques recueillies aux termes de CIM antérieures?

   Cette question ne relève pas d'une analyse quantitative. Cependant, à cause des renvois que comportent un grand nombre de codes, la réponse à cette question devra se faire par une analyse informatique. Il faudra mettre sous forme ordinolingue les tables de conversion des CIM antérieures et établir les critères de comparaison dans les deux sens.

   Si la comparaison établit un lien "un à un" entre les codes, la spécificité des nouveaux codes sera alors la même que celle des anciens et toutes les comparaisons seront possibles. Cependant, il est peu probable que les codes correspondent parfaitement les uns aux autres. Les nouveaux pourront être plus précis que les anciens dans une catégorie de maladies, auquel cas un ancien code pourra correspondre à plus d'un nouveau code. L'inverse pourrait également se produire, c'est-à-dire un nouveau code correspondant à plusieurs anciens codes. Ces conditions imposent des limites au regroupement des codes. L'analyse des groupes de codes interreliés pourra cerner les problèmes de comparaison.

3. La CIM facilite-t-elle l'agrégation de codes en catégories valables?

   Cette question devra être étudiée avec l'aide de spécialistes et d'utilisateurs. On pourra demander aux répondants de proposer des catégories valables ou de regrouper des schèmes; ceux-ci pourront correspondre à leurs propres catégories ou à des catégories utilisées dans leur spécialité. On pourra alors étudier, de façon informelle et par des programmes de regroupement, la facilité et la possibilité de regrouper les catégories de la CIM pour qu'elles correspondent à ces catégories. A l'inverse, les catégories qui découleraient de l'omission pure et simple des

derniers chiffres du code devraient être étudiées par des spécialistes pour qu'ils nous indiquent si ces schèmes de regroupement mènent à des catégories valables.

4. Les analystes peuvent-ils établir facilement les codes qui correspondent à un diagostic donné?

   La réponse à cette question reposera dans une large mesure sur la qualité de la documentation, les index et les renvois de la CIM ainsi que sur sa structure logique. Il faudra également prendre en compte l'opinion des personnes qui ont participé à la révision de la CIM et les documents de référence qui l'accompagnent. En outre, on devra remettre aux analystes la documentation et leur demander s'il leur a été facile de trouver les codes qui correspondent à des diagnostics donnés. Leur tâche serait de comparer la CIM à la version précédente, plus particulièrement les points forts et les faiblesses de la documentation et les répercussions de la structure logique de la classification sur la facilité de recherche des codes.

5. L'analyse est-elle possible dans une diversité de milieux?

   Il y aura lieu de demander aux divers organismes et utilisateurs de la CIM si quelque caractéristique du nouveau code empêche son utilisation dans leur milieu; par exemple, ce pourrait être dû à des caractères spéciaux, à des caractères alphabétiques, à la longueur des codes, à la variabilité de la longueur des codes.

6. La CIM donne-t-elle des résultats similaires lorsqu'elle est appliquée dans des milieux différents?
7. Les codeurs peuvent-ils être facilement formés?
8. Quelle est l'efficacité du système de codage?

   L'examen statistique de ces questions devra se faire à l'aide d'expériences soigneusement planifiées et exécutées. Pour toutes ces questions, il faudra comparer la CIM à une version antérieure et solliciter la participation d'un groupe de codeurs ne possédant que peu ou pas d'expérience avec l'une ou l'autre des versions. Quelques codeurs recevront une formation pour travailler sur les deux versions et l'on comparera les résultats du codage.

   Dans le cas de la question 6, la comparaison des deux versions n'est pas aussi importante que la comparaison des codeurs d'antécédents différents (médecins, codeurs formés, déclarants non spécialisés) et la comparaison de différents niveaux de n'importe quel autre facteur environnemental qui pourrait influer sur les déclarations (hôpital privé ou public, etc.)

   Dans le cas de la question 7, nous sommes intéressés de connaître la durée de la période d'apprentissage des codeurs de façon à faire progresser les deux programmes de formation le plus possible en même temps et à évaluer les codeurs à diverses étapes du programme.

   Pour la question 8, il faudrait puiser à diverses sources de diagnostic pour comparer non seulement l'efficacité relative des deux versions mais également pour cherber à connaître quelle source de diagnostic est la plus rentable.

Pour répondre à ces trois questions, nous utiliserons comme bas les diagnostics que doivent coder les agents. Si l'on adopte la méthode de repétition, chaque codeur codera la même donnée deux fois mais en vertu d'une version différente de la CIM. Si le risque de "contamination" est trop élevé, on utilisera un modèle d'interpénétration dans lequel une moitié des agents choisis au hasard utilisera une version et l'autre moitié la deuxième version. L'objectif ultime sera de mesurer les répercussions du système de classification sur le codage.

Existe-t-il une différence importante entre les deux systèmes et, dans l'affirmative, quelle est-elle? Il faudra élaborer soigneusement ces expériences puisqu'il est peu probable qu'un grand nombre de codeurs participeront à l'étude.

9. Le système de codage comporte-t-il des codes pour tenir compt des données déclarées par les non-spécialistes?

   Cette question devra être étudiée par le concepteur de la CIM et par les organismes qui font appel à des non-spécialistes pour recueillir et coder les données en vertu de la CIM.

10. Le système de codage comporte-t-il des codes qui font ressortir les préoccupations des pays en développement?

    Voilà une autre question à laquelle devront répondre les organismes et les utilisateurs en cause. Les codes devront être compatibles avec les besoins des pays utilisateurs en plus de se prêter à des comparaisons avec les données basées sur des codes plus détaillés de CIM.

11. La CIM permet-elle l'ajout de nouveaux codes?

    Seuls les concepteurs de la CIM pourront répondre à cette question. Il faudrait demander aux experts de nous indiquer les domaines les plus susceptibles de faire l'objet de nouveaux codes. Il faudra ensuite vérifier s'il est réellement possible d'ajouter de nouveaux codes dans les sections correspondantes de la CIM.

12. La CIM intégrera-t-elle des nosographies pour des utilisations variées?

    Oui, seulement si les divers utilisateurs sont convaincus de l'utilité de la chose. Les concepteurs de la CIM devraient élaborer une classification connexe pour des usages spéciaux, ou inciter d'autres personnes à le faire, et s'assurer que les raisons motivant d'autres utilisateurs à établir un lien entre leurs nosographies et la CIM justifient pleinement la démarche.

13. Coût de recyclage des codeurs.

    Les spécialistes de la formation deraient êtr en mesure d'évaluer ce coût. On peut obtenir des chiffres plus précis en contrôlant le processus réel de recyclage. Cependant, on recommande de ne pas procéder de cette façon puisqu'il n'est pas possible de mesurer les coûts non matériels et que de toute façon, même l'évaluation des coûts directs s'avère une tâche difficile.

14. Coût des nouveaux logiciels informatiques.

Une CIM révisée nécessitera un changement des logiciels informatiques pour l'extraction, le stockage, l'analyse et la diffusion. On pense que le coût dépendra de la nature du logiciel dans une installation donnée. Cependant, il devrait être possible d'évaluer de façon générale les répercussions d'une nouvelle CIM sur le coût d'achat de nouveaux logiciels. En outre, on pourrait étudier plusieurs exemples de systèmes informatiques et décider ensuite comment ils pourraient être modifiés.

15. Coût du passage des programmes informatiques.

On recommande une approche similaire à celle donnée à la question 14, c'est-à-dire une évaluation générale accompagnée d'une analyse de cas particuliers. Le point important ici c'est la performance (c'est-à-dire la demande de temps d'ordinateur et d'espace de stockage) du logiciel fonctionnant dans la nouvelle version de la CIM.

16. Coût de reconception des procédures de diffusion.

Les experts techniques, qui sont familiers avec les procédures de diffusion utilisées pour la CIM-9, devront analyser dans quelle mesure la nouvelle CIM influera sur les procédures de diffusion. Parmi les questions possibles à l'étude, mentionnons la structure des tableaux, la présentation des rapports et les méthodes d'analyse et de présentation des données.

17. Coût des comparaisons avec les données basées sur la CIM-9.

Il est possible (et recommandé) d'évaluer simultanément un certain nombre de solutions de rechange à une CIM révisée. De cette façon, il serait possible de réduire les coûts de l'évaluation et de mettre à l'essai un grand nombre de révisions possibles.